



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

경제학석사학위논문

출산과 여성 노동공급
출산력을 이용한 도구변수 추정

2018년 2월

서울대학교 대학원
경제학부 경제학 전공

김 규 식

출산과 여성 노동공급

출산력을 이용한 도구변수 추정

지도교수 이 철 인

이 논문을 경제학석사 학위논문으로 제출함.

2017년 10월

서울대학교 대학원

경제학부 경제학 전공

김 규 식

김규식의 경제학석사 학위논문을 인준함.

2017년 12월

위 원 장 김 봉 근 (인)

부 위 원 장 이 철 인 (인)

위 원 이 정 민 (인)

출산과 여성 노동공급

출산력을 이용한 도구변수 추정

김규식

서울대학교 경제학부

2018년 2월

국문초록

본 논문에서는 자녀의 숫자가 여성의 노동공급에 어떻게 영향을 미치는지 분석한다. 의학계에서 임의성을 가진다고 알려진 첫번째 임신에서의 유산을 도구변수로 사용하여 자녀의 수가 여성의 노동공급에 미치는 영향을 추정하였다. 미국의 National Survey of Family Growth의 최신 자료를 이용해 분석한 결과 여성의 노동공급은 자녀의 수가 증가해도 대체로 변하지 않는 것으로 나타났다.

주제어: 출산, 노동공급, 자연실험, 유산, 낙태, 도구변수 추정법 .

학번: 2012-20148.

목차

1. 서론	1
2. 데이터	5
3. 추정전략과 기본결과	9
3.1. 추정전략: 2SLS 추정량	10
3.2. 2SLS 추정치	10
4. 확장된 결과	16
4.1. 선행연구와 표본	17
4.2. 가계의 자원	22
5. 결론	26
참고문헌	28
A. 부록	30
A.1. Hotz et al. (2005)의 추정전략	30
A.2. 기본결과	31
A.3. 선행연구와 표본	32
A.4. 가계의 자원	35

그림 목록

1. 미국의 여성 경제활동 참가율과 자녀의 수	1
2. 도구변수의 설정	6

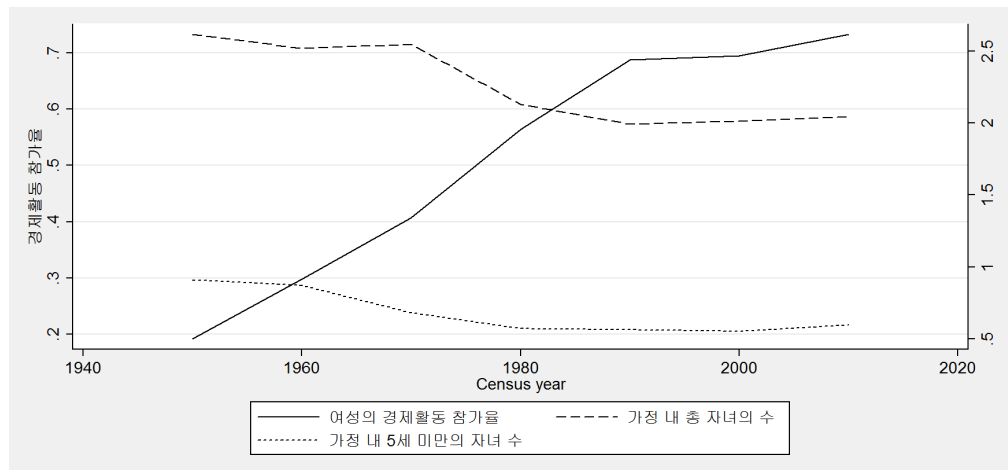
표 목록

1. 표본의 요약통계량: 평균과 표준편차	7
2. 통제집단과 처치집단의 기초변수 차이	8
3. OLS vs. IV 추정치 비교: 경제활동참가	12
4. OLS vs. IV 추정치 비교: 경제활동참가, 전업직장, 모든직장	14
5. 선행연구 계수와의 비교	15
6. 선행연구와 비교 가능한 자녀의 출산 차이: 경제활동참가	19
7. 선행연구와 비교 가능한 자녀의 출산 차이: 전업직장	20
8. 선행연구와 비교 가능한 자녀의 출산 차이: 모든직장	21
9. 가계의 자원 I - 기혼여성과 미혼여성 차이	23
10. 가계의 자원 II - 대리변수로서의 남편의 학력	25
11. OLS vs. IV 추정치 비교: 경제활동참가	33
12. OLS vs. IV 추정치 비교: 경제활동참가, 전업직장, 모든직장	34
13. 선행연구와 비교 가능한 자녀의 출산 차이: 경제활동참가	36
14. 가계의 자원 I - 기혼여성과 미혼여성 차이	37
15. 가계의 자원 II - 대리변수로서의 남편의 학력	38

1. 서론

미국의 여성 경제활동 참가율은 1890년에는 단지 5%에 불과했지만 1980년대에는 50%에 이르렀다 (Goldin, 1991). 2015년 현재는 57.6% 으로 비약적으로 상승했다. 반면 여성이 평생 동안 출생하는 자녀의 숫자는 1960년대에는 3명을 상회했지만 2013년에는 단지 2명 수준에 머무르는데 그치고 있다. 아래의 그림 1은 이러한 모습을 잘 보여주고 있다.

Figure 1: 미국의 여성 경제활동 참가율과 자녀의 수



주석: 위 그림은 미국의 Integrated Public Use Microdata Series 1950년부터 2013년까지의 데이터를 이용한 저자의 계산에 의한 것임을 밝힘

이러한 관찰에 근거하여 경제학자들은 여성의 노동공급과 자녀의 출산 사이의 관계를 연구하였다.

그런데 여성의 노동공급과 자녀의 출산은 동시에 결정된다. 예컨대 자녀의 출산에 수반되는 양육은 여성이 직장생활에 투입할 수 있는 가용시간을 제한하게 된다. 그런데 직장경력에 대한 선호가 큰 여성은 자녀의 출산을 미룰 수 있기 때문에 OLS 추정치는 자녀의 출산이 노동공급에 미치는 효과를 과소 또는 과대평가할 가능성이

있다.

본 논문에서는 이러한 내생성 문제를 해결할 수 있는 일치추정량을 구하기 위해 Hotz et al. (2005)와 Li (2009)를 따라 첫번째 임신에서의 유산 여부를 도구변수로 사용한다. 의학계에 의하면 유산은 수정 단계에서의 태아의 염색체가 비정상적으로 형성됨에 따라서 임의적으로 발생한다. 이러한 인지 가능한 유산은 총임신의 대략 12 ~ 15%를 차지하고 있으며, 첫번째 임신에서의 유산여부는 사회경제적인 변수들로 설명하기 어려운데 (Regan and Rai, 2000), 이를 통해 유산의 외생성을 추측해볼 수 있다.

선행연구에서 사용된 주요 도구변수들과 비교하여 첫번째 임신에서의 유산 여부는 몇 가지 장점을 가지기 때문에 이를 사용하였다. 선행연구에서 사용되는 주요 도구변수로는 쌍둥이 출산, 형제의 성별조합, 그리고 첫번째 시험관 아기 시술에서의 성공여부 등이 있다. 아래에서는 해당 도구변수를 이용한 연구들의 성과와 한계를 살펴본다. 그리고 본 논문이 우리가 사용하는 도구변수를 이용한 선행연구를 어떻게 계승하고 발전시키는지에 대해 논의한다.

먼저 쌍둥이 출산 도구변수를 살펴보자. 쌍둥이 출산은 임의적으로 발생하는 사건이라고 생각되므로 도구변수로써 사용되었다. 이를 사용한 연구는 Rosenzweig and Wolpin (1980), Bronars and Grogger (1994), 그리고 Angrist and Evans (1998)을 들 수 있다. 그런데 미국에서 비히스패닉계 백인 여성의 경우 1980년대부터 2000년대에 이르는 30년 동안 쌍둥이 또는 다생아 출생비율이 두 배로 증가하는 현상이 관찰된다 (Martin et al., 2012). 이러한 다생아 비율의 증가는 불임 치료를 위한 배란촉진제의 투약에 의한 것으로 의심된다 (Schultz, 2007). 이때 쌍둥이 또는 다생아의 출산은 여성의 불임치료 선택 여부에 영향을 받을 수 있기 때문에 임의적으로 발생한다고 말하기 어려울 수 있다.

다음으로 형제의 성별조합 도구변수를 살펴보자. 태아 성별의 결정은 임의적이고

부모는 자녀의 성별이 고른 것을 선호하기 때문에 처음 낳은 두 명의 자녀의 성별이 동일할 때 출산을 한 번 더 함으로써 성별을 고르게 하려고 한다는 사실에 근거하여 형제의 성별조합이 도구변수로 사용되고 있다. 이를 이용한 연구는 Angrist and Evans (1998), Cruces and Galiani (2007), Hirvonen (2009), Angelov and Karimi (2012) 등을 들 수 있다.

그런데 형제의 성별조합 도구변수를 적용할 때의 표본은 적어도 자녀를 2명 이상 가지고 있는 여성으로 제한된다. Angrist and Evans (1998)는 미국의 경우 출산 감소가 과거에는 자녀를 3명 가지던 것이 2명 가지는 것으로 줄어든 현상에서 발생하므로, 성별조합 도구변수를 사용해도 대표성에 문제가 없다고 한다. 우리의 도구변수인 첫번째 임신에서의 유산 여부는 한 번 이상 임신을 한 여성이 표본이므로 형제의 성별조합 도구변수와 같이 대표성을 가지는데, 형제 성별조합 도구변수와 달리 저출산으로 인해 자녀의 수가 2명 이하인 국가에서도 표본이 대표성을 가질 수 있다.

마지막으로 첫번째 시험관 아기 시술에서의 성공여부를 살펴보자. Lundborg et al. (2014)는 첫번째 시험관 아기 시술의 성공여부는 임의적으로 결정된다는 판단 하에 자녀의 출산이 노동성과에 미치는 영향을 분석하였다. 하지만 저자들도 인정하듯이 분석 대상이 되는 표본이 여성 전체를 대표하기는 어려워 보인다. 시험관 아기 시술을 받은 집단의 여성들은 모집단 전체의 여성에 비해서 나이가 많고, 교육을 더 받았으며, 소득도 더 많기 때문이다.

반면 우리의 도구변수인 첫번째 임신에서의 유산 여부는 적어도 한 번 이상 임신을 한 여성을 표본으로 삼는다. 비록 임신을 하는 여성들은 출산을 고려하지 않는 여성을 포함하는 여성 전체와는 다를 수 있지만, 출산을 고려하는 여성 전체를 표본으로 하기 때문에 출산하고자 하는 여성을 대표할 수 있을 것이다.

이제 본 논문이 동일한 도구변수를 사용하는 선행연구들을 어떻게 계승하고 받

전시키는데 대해 살펴보자. Hotz et al. (2005)는 National Longitudinal Suevey의 NLSY79 데이터에서 해당 도구변수를 이용하여 십대 여성의 출산이 교육, 소득, 노동공급 등 여성의 생애 전반에 어떤 영향을 미치는지를 살펴보았다. 저자들은 십대에 출산을 한 여성과 그렇지 않은 여성 사이의 평균인과효과를 살펴보았는데, 수학적 도출을 통하여 그것이 유산을 겪지 않은 여성과 유산을 겪은 여성의 종속 변수의 차이를 유산을 겪지 않은 여성의 비율로 나눈 것임을 보였다. 그리고 Li (2009)는 National Survey of Family Growth 1988, 1995, 2002를 이용하여 출산이 노동시장에서의 결과에 미치는 영향을 살펴보았다. 하지만 Li (2009)는 20년 전의 과거 데이터를 사용함으로써 현재의 경제에 주는 시사점이 적으며, 사산이나 자궁외임신 등은 실질적으로 유산과 같은 결과가 나타남에도 불구하고 데이터상에 유산이라고 표시된 것만을 고려하였다.

본 논문은 Hotz et al. (2005)을 따라 첫번째 임신에서의 유산을 도구변수로 삼고 Li (2009)을 보강하기 위해 다음과 같은 작업을 수행하였다. 먼저 데이터를 National Survey of Family Growth 2006 – 2015를 분석함으로써 비교적 최근의 노동시장을 다룬다. 둘째, 데이터 상의 유산 뿐만 아니라 사산과 자궁외임신을 광의의 유산으로 포괄하여 통제집단으로 포함시킴으로써 데이터의 배제를 최소화하고자 하였다. 셋째, 낙태는 도구변수의 임의성을 방해할 수 있기 때문에 낙태를 표본에서 제외하였다. 마지막으로, 도구변수가 적용된 집단과 그렇지 않은 집단 사이에 사회경제적 변수에 차이가 있는지 확인하였다. 이는 도구변수의 외생성을 확인하기 위해 필요한 작업이지만 Li (2009)에서는 수행되지 않았기 때문에 추가하였다.

본 논문의 나머지는 다음과 같이 구성된다. 2절에서는 본문에서 사용된 데이터에 대해 설명을 설명하고 첫번째 임신에서 출산한 여성과 유산한 여성 사이의 기초변수의 차이를 살펴본다. 3절에서는 추정전략을 소개하고 기본적인 추정결과를 제시한다. 4절에서는 다양한 각도에서 추정전략을 적용하도록 한다. 5절에서는

분석결과를 정리하도록 한다.

2. 데이터

본 논문에서는 미국의 Centers for Disease Control and Prevention에서 수집하고 발표하는 National Survey of Family Growth 2006 – 2015 (이하 NSFG)를 사용한다. NSFG는 반복횡단면자료로써 여성 23,579명에 대한 데이터를 가지고 있다. 수집된 세부항목으로는 출산력, 불임, 피임도구, 건강상태, 가족의 생활, 그리고 결혼 및 이혼에 대한 다양한 정보가 있다.

NSFG의 특기할 만한 사항은 개별 여성이 경험한 각각의 임신과 그 결과에 대한 상세한 자료를 수집한다는 것이다. 예컨대 첫번째 임신의 결과가 출산, 낙태, 유산, 사산, 자궁외임신, 또는 현재 임신중인지를 기록한다. 그리고 만약 출산으로 끝났다면 그때 태어난 자녀의 숫자는 몇 명인지, 각각의 자녀의 성별은 무엇인지를 기록한다.

하지만 NSFG가 제공하는 경제변수는 풍부하지 않다. 아래에서 살펴보게 될 경제활동참가를 제외한 여타 경제변수들은 없다고 할 수 있다. 예컨대 소득이나 노동시간에 대한 자료는 수집되지 않았다.

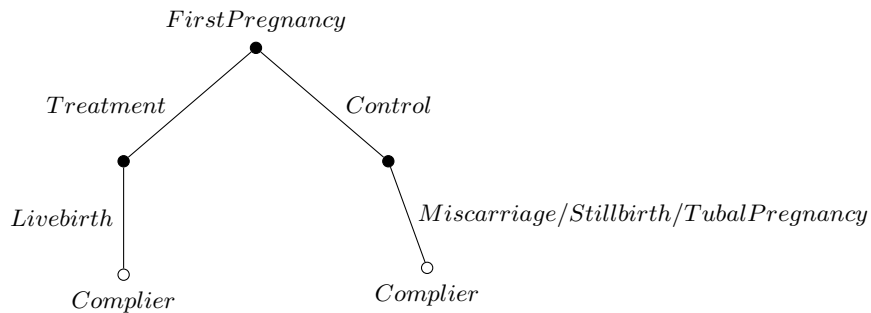
분석의 핵심이 되는 도구변수는 다음과 같이 만들어졌다. 여성의 첫번째 임신의 결과는 출산, 낙태, 유산, 사산, 자궁외임신, 그리고 현재 임신중으로 나눌 수 있다. 이 중에서 현재 임신중인 경우는 표본에서 제외하였다. 이 경우는 아직 결과가 확정되지 않았기 때문이다.

먼저 출산의 경우는 자녀의 수가 증가하게 되므로 도구변수는 1의 값을 가지는 처치집단에 속하게 된다. 한편 낙태의 경우는 행위 주체자의 선택에 의하여 결정된 것이다. 즉, 낙태된 태아는 그러지 않았다면 출산될 수도 있었으며 유산될 수도

있는 것이다. 따라서 이러한 행위는 유산의 임의성을 방해하므로 우리의 표본에서 제외하였다.¹

다음으로 유산, 사산, 자궁외임신은 모두 도구변수가 0의 값을 가져 자녀의 숫자가 증가하지 않는 통제집단으로 설정하였다. 세 가지 경우 모두 자녀의 수가 증가하지 않는 방향으로 작용하기 때문이다. 아래의 그림 2는 이러한 변수설정을 요약해서 보여주고 있다.

Figure 2: 도구변수의 설정



아래의 표 1에는 우리가 다루는 표본의 요약통계량이 제시되어 있다. 전체표본 12,382명의 경제활동참가율은 69%로 전체평균 57%보다 높다. 이는 우리의 표본이 전체 여성과 다르기 때문일 수 있다. BLS가 발표하는 여성의 경제활동참가율은 16세 이상의 모든 여성을 표본으로 하는 반면, NSFSG는 15 ~ 49세의 여성만을 조사하고 있다. 그런데 BLS에 의하면 55세 이상의 여성의 경제활동참가율은 34.7%로 떨어지는 것으로 보고되고 있다. 따라서 이러한 경제활동참가율의 차이가 표본의 차이에서 왔다고 생각해볼 수 있다.

자녀의 수는 2.07명으로 IPUM USA에서 보고되는 것처럼 2명 내외의 수준을 보여주고 있다. 교육연수는 12.90년으로 고등학교 졸업이 평균인 것으로 나타났다.

¹Hotz et al. (2005)은 낙태를 포함한 추정을 하였다. 이러한 표본의 차이에 따른 결과는 부록에서 확인하도록 한다.

연령은 설문 당시에 32.25세였으며, 첫 임신시의 연령은 21.76세인 것으로 나타난다. 그리고 흑인인 여성은 표본의 23%, 히스패닉인 여성은 표본의 27%인 것으로 나타나고 있다. 마지막으로 도시거주자인 여성은 40%인 것으로 나타난다.

Table 1: 표본의 요약통계량: 평균과 표준편차

	평균	표준편차
경제활동참가율	0.69	0.46
자녀의 수	2.07	1.24
혼인여부	0.46	0.50
교육연수	12.90	2.63
연령	32.25	6.95
첫임신 시 연령	21.76	5.20
인종: 흑인 여부	0.23	0.42
인종: 히스패닉 여부	0.27	0.44
도시거주 여부	0.40	0.49
총관측치 수	12,382	

이제 도구변수가 적절한지를 확인해보도록 하자. 도구변수가 적절하기 위해서는 우리의 관심변수인 경제활동참가율에 영향을 주는 경로가 오직 자녀의 출산에 따라서 이루어져야 한다. 따라서 도구변수는 자녀의 출산에만 차이만을 만들 뿐 다른 배경변수들에 영향을 주지 않아야 한다. 아래의 표 2는 각 집단별로 배경변수들을 비교할 수 있는 자료가 된다.

표 2의 계수는 해당변수에 대한 각 집단의 평균을 나타내고 괄호 안의 숫자는 표준편차를 나타낸다. 유산한 여성은 2,225 명으로 전체집단의 17.97% 를 구성한다. 이는 총임신의 12 ~ 15 %가 유산으로 끝난다는 의학계의 보고보다 조금 큰 수치이다.

경제활동 참가율은 유산한 여성은 70.4%, 출산한 여성은 68.9%지만 p-값은 0.15로 두 집단은 유의수준 0.05에서 차이가 없는 것으로 보인다. 한편 우리가 관심을

Table 2: 통제집단과 처치집단의 기초변수 차이

	유산한 여성	출산한 여성	전체	p-값
경제활동참가율	0.704 (0.456)	0.689 (0.463)	0.691 (0.462)	0.15
설문시점까지 출산한 자녀의 수	1.41 (1.257)	2.22 (1.194)	2.07 (1.245)	0.00
교육연수	13.21 (2.636)	12.84 (2.621)	12.90 (2.627)	0.00
연령	32.07 (7.291)	32.29 (6.872)	32.25 (6.949)	0.173
임신 당시의 연령	22.07 (5.603)	21.69 (5.104)	21.76 (5.199)	0.00
혼인여부	0.442 (0.497)	0.458 (0.498)	0.455 (0.498)	0.167
임신 당시의 혼인여부	0.290 (0.454)	0.322 (0.467)	0.316 (0.465)	0.00
인종: 흑인 여부	0.240 (0.427)	0.226 (0.418)	0.228 (0.420)	0.17
인종: 히스패닉 여부	0.212 (0.409)	0.284 (0.451)	0.271 (0.445)	0.00
도시 거주 여부	0.396 (0.489)	0.397 (0.489)	0.397 (0.489)	0.90
관측치 수	2,225	10,157	12,382	

주석: 계수는 각 집단의 평균값이며 괄호 안의 숫자는 각 집단의 표준편차임. p-값은 두 집단의 평균에 차이가 없다는 귀무가설에 대해 계산되었음.

가지는 변수인 설문시점까지 출산한 자녀의 숫자는 유산한 여성은 약 1.41명인 반면 출산한 여성은 약 2.21명인 것으로 나타났다. p-값으로 확인할 수 있듯이 자녀의 수는 첫번째 임신에서 출산한 여성이 유산한 여성에 비해서 더 많다는 것을 확인할 수 있다.

교육연수는 유산한 여성이 13.21년인 반면 출산한 여성은 12.84년으로 약 0.4년의 유의한 차이가 난다. 하지만 그 절대적 차이는 크지 않은 것으로 보인다. 한편 설문시점의 혼인여부는 두 집단 사이에 차이가 없지만 임신 당시의 혼인여부는 차이가 있는 것으로 나타난다. 또한 설문시점의 연령에는 차이가 없었지만 임신시점의 연령에는 차이가 있는 것으로 보인다. 약 0.4세의 차이가 있으며 유산한 여성이 더 어린 것으로 나타난다. 다만 우리가 임신시점이 아니라 임신결과를 기준으로 연령을 측정할 경우에는 두 집단 사이에 차이가 없어지는 것을 확인할 수 있었다. 인종 중 흑인의 비율은 두 집단이 동일하지만 히스패닉의 비율은 차이가 나는 것으로 보인다. 마지막으로 도시에 거주하는 비율은 두 집단 사이에 차이가 없는 것으로 나타난다.

우리는 두 집단 사이에 임신시점의 연령과 교육연수의 차이가 있음을 보았으나 그 차이가 크지 않다는 것을 확인할 수 있었다. 한편 첫번째 임신에서 출산한 여성과 유산한 여성 사이에 우리의 관심변수인 자녀의 숫자가 유의미한 차이를 보이는 것을 보았다. 아래에서는 도구변수가 적절하게 선택되었다는 가정하에, 이를 이용하여 자녀의 수가 여성의 노동공급에 미치는 영향을 추정하도록 한다.

3. 추정전략과 기본결과

3절에서는 추정전략을 소개하고 이에 따른 기본적인 결과들을 제시하도록 한다. Angrist and Pischke (2009)를 따라 2SLS 추정량과 그에 따른 결과들이 제시될

것이다.

3.1. 추정전략: 2SLS 추정량

$$C_i = Z_i\gamma_1 + X_i\gamma_2 + \epsilon_i \quad (1)$$

$$Y_i = C_i\beta_1 + X_i\beta_2 + \eta_i \quad (2)$$

우리의 회귀모형에서 Y_i 는 경제활동 참가여부 등 관심 노동공급 종속변수이며 C_i 는 자녀의 숫자이다. 그리고 우리의 도구변수가 Z_i 일 때, 우리의 관심변수인 β_1 는 회귀식 (1)로부터 얻은 예측치로 관심 인과회귀식 (2)의 C_i 를 대체한 뒤 회귀분석을 함으로써 얻을 수 있다는 것이 잘 알려져있다.

외생 공변량은 교육연수, 연령, 연령의 제곱, 설문시점의 연도더미들, 인종변수, 그리고 도시거주 여부가 될 것이다. 그리고 도구변수는 첫번째 임신에서 자연에 의해 어느 집단에 속하게 되었는지와 그와 밀접하게 관련이 있는 임신결과가 나오기까지의 임신 길이 더미이다. 3.2절에서는 이를 이용한 기본 추정치들을 확인해 본다.

3.2. 2SLS 추정치

먼저 자녀의 숫자가 여성의 경제활동참가에 미치는 효과를 살펴보자. 이는 Angrist and Evans (1998)에서 분석되었듯이 출산을 얼마나 했는지가 노동공급에 미치는 효과를 나타낸다. 아래의 표 3은 이를 보여준다. 먼저 핵심변수들 이외에는 연도 더미만을 통제한 첫번째 열과 두번째 열을 비교해보자. 첫번째 열은 OLS 추정치를, 두번째 열은 IV 추정치를 보여주고 있다. OLS 추정치는 0.01의 유의수준에서 약 4.33%의 경제활동참가 하락을 예측하는 반면에 IV 추정치는 0.05의 유의수준에서 약 3.07%의 하락을 예측하고 있다. OLS 추정치의 경우에 직장생활이나 경력

등 노동시장에 대한 애착이 큰 여성일수록, 노동공급을 늘리는 것보다 자녀를 더 적게 낳기 때문에 계수의 크기가 과대추정될 수 있다. 하지만 도구변수추정의 사후검정을 나타내는 마지막 행들을 살펴볼 때 점추정치를 조심스럽게 받아들여야 한다.

두번째 열에서, 1차 회귀분석의 F-값은 18.52로 기준점인 10을 넘는다. 하지만 Wu-Hausman 검정 결과 우리의 내생변수가 외생적이라는 귀무가설을 기각할 수 없다. 또한 Sargan 검정 결과 우리의 도구변수는 오차항과 상관이 있는 것으로 나타난다. 이러한 결과로 미루어 볼 때 도구변수 추정을 할 필요가 없는 것 처럼 보인다.

그런데 이러한 결과는 다른 통제변수를 추가하면 달라진다. 교육연수, 연령과 그 제곱, 임신 당시의 연령과 그 제곱을 추가적으로 통제한 세번째 열과 네번째 열에서는 OLS 추정치는 다소 커졌지만 IV 추정치는 0.1의 유의수준에서 유의하지 않은 것을 확인할 수 있다. 인종과 도시거주 여부를 추가로 통제한 다섯번째 열과 여섯번째 열에서도 여전히 IV 추정치는 유의하지 않다.

그런데 통제변수를 추가했을 때는 1차 회귀분석의 F-값이 20을 넘고, Wu-Hausman 검정 결과 우리의 내생변수가 외생적이라는 귀무가설을 기각할 수 있으며, Sargan 검정 결과 우리의 도구변수는 오차항과 상관이 없는 것으로 나타나는 것을 알 수 있다. 우리의 결과에 비추어 볼 때 OLS 추정치는 자녀의 숫자가 여성의 경제활동 참가에 미치는 효과를 과대추정하며, IV 추정치에 의하면 여성의 경제활동참가에 자녀의 숫자가 영향을 미치지 않는 것으로 나타난다.

표 3에서는 경제활동참가 확률이 자녀의 숫자에 따라 달라지는지를 살펴보았고 계수가 유의하지 않은 것을 보았다. 그런데 노동공급을 직장의 형태에 따라 전업직장 여부와 어떠한 직장이라도 가지고 있는지 여부로 나누었을 경우에 다른 결론이 나올 수 있으므로 이를 추가적인 노동공급 변수로서 살펴본다.

Table 3: OLS vs. IV 추정치 비교: 경제활동참가

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
자녀의 수	-0.0435*** (0.00341)	-0.0307** (0.0127)	-0.0575*** (0.00398)	-0.00447 (0.0141)	-0.0570*** (0.00398)	-0.00379 (0.0141)
교육연수			0.141*** (0.0134)	0.160*** (0.0144)	0.133*** (0.0134)	0.151*** (0.0145)
교육연수의 제공			-0.00405*** (0.000484)	-0.00468*** (0.000528)	-0.00377*** (0.000487)	-0.00438*** (0.000529)
연령			0.0514*** (0.00566)	0.0304*** (0.00774)	0.0514*** (0.00566)	0.0304*** (0.00773)
연령의 제공			-0.000649*** (0.0000861)	-0.000387*** (0.000109)	-0.000648*** (0.0000861)	-0.000386*** (0.000109)
임신 당시 연령			-0.0180*** (0.00606)	-0.00684 (0.00676)	-0.0151** (0.00608)	-0.00383 (0.00679)
임신 당시 연령의 제공			0.000181 (0.000122)	0.0000634 (0.000129)	0.000137 (0.000122)	0.0000172 (0.000129)
인종: 흑인					0.0540*** (0.00961)	0.0573*** (0.00998)
도시 거주 여부					-0.00118 (0.00845)	-0.00273 (0.00846)
연도 더미	통제	통제	통제	통제	통제	통제
1차 회귀분석의 F 값		18.52		20.17		20.16
Wu-Hausman 검정 p-값		0.295		0.0000830		0.0000771
Sargan 검정 p-값		0.00652		0.292		0.221
총 관측치 수	12382	12382	12382	12382	12382	12382

주석: 표준오차 값은 괄호 안에 표시되어 있다. 종속변수는 경제활동참가 여부이다. 홀수 열은 OLS 추정의 결과이며 짝수 열은 IV 추정의 결과로 대칭적으로 비교할 수 있다.

표 4 에서 홀수 열은 OLS 추정치이며 짝수 열은 IV 추정치이다. 그리고 (1), (2)열은 경제활동참여, (3), (4)열은 전업직장, (5), (6)열은 모든 직장에 대한 추정치들이다. 짝수 열의 가장 하단 패널에는 도구변수추정의 사후검정의 통계량인 1차 회귀분석의 F-값, Wu-Hausman 검정의 p-값, 그리고 Sargan 검정의 p-값이 제시되어 있다.

전업직장의 경우 자녀의 숫자에 따라 0.90%의 노동공급 감소가 예측되지만 유의수준 0.1에서 유의하지 않은 것으로 나타났다. 하지만 1차 회귀분석의 F-값은 20.16으로 10을 넘는다. 또한 Wu-Hausman 검정의 p-값이 0에 가까우므로 자녀의 수는 외생변수라는 귀무가설을 기각할 수 있다. 마지막으로 Sargan 검정의 p-값을 통해 우리의 도구변수는 오차항과 상관이 없는 것으로 나타나는 것을 알 수 있다.

마찬가지로 모든직장의 경우 자녀의 숫자에 따라 0.48%의 노동공급 감소가 예측되지만 유의수준 0.1에서 유의하지 않은 것으로 나타났다. 하지만 1차 회귀분석의 F-값은 20.16으로 10을 넘는다. 또한 Wu-Hausman 검정의 p-값이 0에 가까우므로 자녀의 수는 외생변수라는 귀무가설을 기각할 수 있다. 마지막으로 Sargan 검정의 p-값을 통해 우리의 도구변수는 오차항과 상관이 없는 것으로 나타나는 것을 알 수 있다.

우리는 표 4의 결과를 통해 자녀의 수가 노동공급에 영향을 미친다고 말할 수 없음을 확인할 수 있었다. 이러한 추정치들을 자녀의 출산이 노동공급에 미치는 영향을 분석한 선행연구들과 비교해볼 수 있다. 먼저 미국의 1980년 PUMS 데이터에서 미혼 여성의 쌍둥이 출생을 도구변수로 이용한 Bronars and Grogger (1994)는 3.5%의 노동공급(노동참가여부) 감소를 보고하였다. 마찬가지로 미국의 1990년 PUMS 데이터에서 형제의 성별조합을 도구변수로 이용한 Angrist and Evans (1998)는 노동공급(모든직장)의 감소를 6.3%로 보고했다. 마찬가지로 NSFG 1988, 1995, 2002를 사용한 Li (2009)는 2.4%의 노동공급(모든직장) 감소를 보고했지만

Table 4: OLS vs. IV 추정치 비교: 경제활동참가, 전업직장, 모든직장

	(1) 경제활동참가	(2)	(3) 전업직장	(4)	(5) 모든직장	(6)
자녀의 수	-0.0570*** (0.00381)	-0.00379 (0.0141)	-0.0568*** (0.00409)	0.00905 (0.0151)	-0.0575*** (0.00397)	0.00485 (0.0147)
교육연수	0.133*** (0.0136)	0.151*** (0.0145)	0.109*** (0.0145)	0.132*** (0.0155)	0.128*** (0.0141)	0.150*** (0.0151)
교육연수의 제곱	-0.00377*** (0.000502)	-0.00438*** (0.000529)	-0.00309*** (0.000538)	-0.00384*** (0.000568)	-0.00345*** (0.000522)	-0.00416*** (0.000551)
연령	0.0514*** (0.00555)	0.0304*** (0.00773)	0.0595*** (0.00594)	0.0336*** (0.00831)	0.0569*** (0.00577)	0.0323*** (0.00806)
연령의 제곱	-0.000648*** (0.0000849)	-0.000386*** (0.000109)	-0.000736*** (0.0000910)	-0.000411*** (0.000117)	-0.000707*** (0.0000884)	-0.000400*** (0.000113)
임신 당시 연령	-0.0151** (0.00611)	-0.00383 (0.00679)	-0.0132** (0.00654)	0.000797 (0.00730)	-0.00945 (0.00635)	0.00378 (0.00708)
임신 당시 연령의 제곱	0.000137 (0.000124)	0.0000172 (0.000129)	0.0000805 (0.000133)	-0.0000682 (0.000139)	0.0000429 (0.000129)	-0.0000980 (0.000134)
인종: 흑인여부	0.0540*** (0.00988)	0.0573*** (0.00998)	0.0344*** (0.0106)	0.0386*** (0.0107)	0.0179* (0.0103)	0.0218** (0.0104)
도시거주 여부	-0.00118 (0.00840)	-0.00273 (0.00846)	-0.0186** (0.00900)	-0.0206** (0.00909)	-0.00998 (0.00873)	-0.0118 (0.00882)
연도 더미	통제	통제	통제	통제	통제	통제
1차 회귀분석의 F 값		20.16		20.16		20.16
Wu-Hausman 검정 p-값		0.0000771		0.00000505		0.00000848
Sargan 검정 p-값		0.221		0.254		0.288
총 관측치 수	12382	12382	12382	12382	12382	12382

주석: 표준오차 값은 괄호 안에 표시되어 있다. 종속변수는 (1), (2)열은 경제활동참가 여부이고 (3), (4)열은 전업직장, (5), (6)열은 모든직장이다. 홀수 열은 OLS 추정의 결과이며 짝수 열은 IV 추정의 결과이다.

계수는 유의하지 않았다. 아래의 표 5에서는 이러한 추정치를 비교하고 있다.

Table 5: 선행연구 계수와의 비교

	노동공급 변수	대상 표본	대상 연도	계수
Bronars and Grogger (1994)	경제활동참가	자녀 1명에서 자녀 2명	1980년	-0.035 (0.017)
Angrist and Evans (1998)	모든직장	자녀 2명에서 자녀 3명	1990년	-0.063 (0.018)
Li (2009)	모든직장	자녀 0명에서 자녀 1명	1988, 1995, 2002 년	-0.024 (0.022)
김규식(2018)	모든직장	자녀 0명에서 자녀 1명	2006 ~ 2015년	-0.004 (0.0147)

이러한 비교는 실상은 무의미한 것일 수 있다. 그 이유는 크게 두 가지이다. 첫째, 각각의 선행연구들이 다루고 있는 표본이 다르다. Bronars and Grogger (1994)는 첫 출산인 여성인 미혼 여성, Angrist and Evans (1998)는 적어도 두 명 이상의 자녀를 가진 여성, 그리고 Li (2009)는 적어도 한 번 이상 임신을 한 여성이 표본이 된다. 즉 각각의 연구들은 차례로 자녀가 1명에서 2명, 2명에서 3명, 그리고 0명에서 1명으로 넘어가는 서로 다른 출산의 차이(margin)가 노동공급에 미치는 영향을 측정하고 있기 때문이다. 만약 이러한 차이가 선행이 아니라면 우리가 관심을 가지는 계수의 크기가 달라질 수 있다.

둘째, 각각의 선행연구들은 다른 연도의 노동시장을 다루고 있기 때문에 매해의 노동환경이 크게 다를 수 있다. 평균적인 경제활동참가율을 기준으로 살펴볼 때, 1980년도는 63.8%, 1990년도는 66.5%, 1988, 1995, 2002년도에는 66.4%, 2006-2015년도에는 64.5%로 작지만 차이가 존재하기 때문이다. 또한 평균적인 실업율은 7.2%, 5.6%, 5.6%, 그리고 7.0%로 약간의 차이가 나는 것을 확인할 수 있다. 하지만 우리는 연도별 노동시장의 변동을 통제하기 위해서 연도 터미를 이미 통제하고 있기 때문에 노동시장에 커다란 단절이 발생하지 않는 이상, 노동시장의 환경이

추정치에 큰 영향을 미치지 않는다고 생각된다.

그럼에도 불구하고 우리의 추정치는 선행연구의 결과들과 달리 그 크기가 작으며 유의하지 않다. 다음 절에서는 이러한 차이가 우리가 다루는 표본과 선행연구들이 다루는 표본이 다르다는 점에 주목하여, 이를 고려한 새로운 추정을 수행하도록 한다.

4. 확장된 결과

앞서 3절에서는 총자녀의 수가 노동공급에 얼마나 영향을 미치는지 살펴보았고 그 결과 노동공급에 변화가 없음을 확인했다. 이번 절에서는 기본 모형을 다음과 같이 확장해서 추정하도록 한다.

먼저 도구변수를 두 번째 또는 세 번째 임신에서 유산을 했는지로 바꾸어본다. 우리는 3절에서 선행연구들과 추정치가 다른 이유가 표본이 달리 설정되었기 때문일 수 있다는 의문을 제기하였다. 우리가 사용하고 있는 도구변수를 두 번째 또는 세 번째 임신에서 유산을 했는지로 확장함으로써 선행연구들이 다루는 표본을 유사하게 만들어낼 수 있다. 이렇게 구한 추정치를 선행연구들과 비교해보도록 한다.

다음으로 여성이 속한 가계의 자원이 다를 경우에 따라 추정치가 달라질 수 있다. 하지만 NSFSG는 가계 총소득만을 제공하기 때문에 직접적인 고려는 어렵다. 따라서 비노동소득이 없는 미혼여성 and 비노동소득이 있는 기혼여성 사이에 자녀의 수가 노동공급에 미치는 영향을 비교해본다. 그리고 표본을 혼인상태에 있는 여성으로 한정하여 아내의 학력과 남편의 학력의 차이를 비노동소득의 대리변수로 삼아 자녀의 수가 노동공급에 미치는 영향을 살펴본다.

4.1. 선행연구와 표본

우리는 3 절의 마지막에 선행연구들의 추정치와 우리가 얻은 추정치의 차이를 유발하는 원인 중의 하나로서 표본이 상이하기 때문이 아닌지 고려해야 한다고 하였다. 이번 절의 목적은 우리의 도구변수를 첫 번째 임신 뿐만 아니라, 두 번째 임신, 세 번째 임신 등으로 확장함으로써 이러한 의심이 맞는지 확인해보고자 함이다..

경제활동참가를 노동공급 변수로 보는 표 6 은 다음과 같이 구성되어 있다. (1), (2)열은 자녀의 수가 0명에서 1명으로 늘어날 수 있는 첫번째 임신을 한 여성들로 구성되어 있다. 이들은 표 3 에서 다룬 표본 그리고 Li (2009)와 동일하다. 그리고 (3), (4)열은 자녀의 수가 1명에서 2명으로 늘어날 수 있는 두 번째 임신을 한 여성들로 구성되어있다. 이들은 Bronars and Grogger (1994)이 다루는 표본에 대응된다. 한편 (5), (6)열은 자녀의 수가 2명에서 3명으로 올라갈 수 있는 세번째 임신을 한 여성들로 구성되어 있다. 이들은 Angrist and Evans (1998)에 대응된다. 마지막으로 각각의 표본에 대하여 홀수 열은 OLS 추정치, 짝수 열은 IV 추정치를 나타낸다.

하단 패널의 도구변수추정의 사후검정을 살펴보면 우리의 도구변수는 잘 선택된 것으로 보인다. 먼저 1차 회귀분석의 F-값은 그 값이 가장 작은 마지막 표본의 경우에도 11.92로 10을 넘는다. 또한 Wu-Hausman 검정의 p-값이 0에 가까우므로 자녀의 수는 외생변수라는 귀무가설을 기각할 수 있다. 마지막으로 Sargan 검정의 p-값을 통해 우리의 도구변수는 오차항과 상관이 없는 것으로 나타나는 것을 알 수 있다.

하지만 OLS 추정치는 일관되게 유의한 경제활동참가 확률의 감소를 보고하는데 반하여, IV 추정치는 모든 집단에서 유의수준 0.1에서도 유의하지 않으면서도 1%도 안 되는 경제활동참가 확률의 감소를 예측하고 있다. 유일하게 Angrist and

Evans (1998)에 대응하는 표본만 2.16%의 경제활동참가 확률의 감소를 비교적 크게 예측하고 있지만, 여전히 유의하지 않은 것을 알 수 있다.

이제 노동공급 변수가 전업직장일 때를 살펴보자. 표 7은 표 6과 마찬가지로 형식으로 구성되어 있다. OLS 추정치는 첫번째 표본에서는 5.68%, 두번째 표본에서는 4.94%, 그리고 세번째 표본에서는 4.00%의 유의미한 전업직장 근무 확률 감소를 예측한다. 하지만 이전의 결과와 마찬가지로 IV 추정치는 세번째 표본을 제외하고는 통계적으로 유의하지 않고 전업직장 근무 확률이 1% 미만으로 감소할 것으로 예측하고 있다. 비록 세번째 표본에서의 IV 추정치인 3.34%의 전업직장 근무확률 감소는 유의하지만, 하단 패널의 도구변수추정의 사후검정을 살펴보면 Wu-Hausman 검정의 p-값에서 볼 수 있듯이 자녀의 수가 외생적이라는 귀무가설을 기각할 수 없다. 이때에는 OLS 추정치가 IV 추정치보다 효율적이므로 IV추정치로 받아들일 수 없어 보인다.

마지막으로 노동공급 변수가 어떠한 직장이든지 일을 하고 있는지 여부를 살펴보자. 표 8은 표 6과 동일하게 구성되어 있다. 여기서도 OLS 추정치는 첫번째 표본에서는 5.75%, 두번째 표본에서는 5.24%, 그리고 세번째 표본에서는 4.30%의 유의미한 모든직장 근무 확률 감소를 예측한다. 하지만 IV 추정치는 세번째 표본을 제외하고는 도구변수추정의 사후검정을 통과하지만 점추정치는 유의하지 않은 것을 확인할 수 있었다.

결국 표 5에서 보았던 우리의 계수가 선행연구들과 다른 것은 표본의 상이함 때문은 아닌 것으로 보인다. 우리의 도구변수를 이용하여 선행연구의 표본에 대응되는 표본을 만들고 분석해보았으나, 모든 집단에서 IV 추정치들은 유의하지 않거나 도구변수가 적절하지 않았기 때문에 자녀의 수가 노동공급에 미치는 효과가 0이라는 귀무가설을 기각할 수 없었기 때문이다.

Table 6: 선행연구와 비교 가능한 자녀의 출산 차이: 경제활동참가

	(1) OLS 0명에서 1명 출산	(2) IV 1명에서 2명 출산	(3) OLS 1명에서 2명 출산	(4) IV 2명에서 3명 출산	(5) OLS 2명에서 3명 출산	(6) IV 2명에서 3명 출산
자녀의 수	-0.0570*** (0.00381)	-0.00379 (0.0141)	-0.0524*** (0.00430)	-0.00736 (0.0157)	-0.0446*** (0.00522)	-0.0216 (0.0178)
교육연수	0.133*** (0.0136)	0.151*** (0.0145)	0.109*** (0.0155)	0.125*** (0.0165)	0.107*** (0.0198)	0.114*** (0.0206)
교육연수의 제곱	-0.00377*** (0.000502)	-0.00438*** (0.000529)	-0.00291*** (0.000578)	-0.00343*** (0.000606)	-0.00281*** (0.000746)	-0.00301*** (0.000761)
연령	0.0514*** (0.00555)	0.0304*** (0.00773)	0.0388*** (0.00737)	0.0237*** (0.00896)	0.0422*** (0.0108)	0.0352*** (0.0120)
연령의 제곱	-0.000648*** (0.0000849)	-0.000386*** (0.000109)	-0.000448*** (0.000110)	-0.000254** (0.000128)	-0.000492*** (0.000159)	-0.000402*** (0.000172)
임신 당시 연령	-0.0151** (0.00611)	-0.00383 (0.00679)	-0.0224*** (0.00748)	-0.0158** (0.00784)	-0.0240** (0.0104)	-0.0226** (0.0104)
임신 당시 연령의 제곱	0.000137 (0.000124)	0.0000172 (0.000129)	0.000268* (0.000158)	0.000197 (0.000161)	0.000289 (0.000228)	0.000286 (0.000228)
인종: 흑인여부	0.0540*** (0.00988)	0.0573*** (0.00998)	0.0519*** (0.0115)	0.0521*** (0.0115)	0.0645*** (0.0149)	0.0630*** (0.0149)
도시거주 여부	-0.00118 (0.00840)	-0.00273 (0.00846)	0.00586 (0.00973)	0.00336 (0.00981)	0.00830 (0.0126)	0.00632 (0.0127)
연도 더미						
1차 회귀의 F 값		20.16		16.11		11.92
Wu-Hausman p-값		0.0000771		0.00266		0.176
Sargan test p-값		0.221		0.462		0.229
총 관측치 수	12382	12382	9653	9653	5919	5919

주석: 표준오차 값은 괄호 안에 표시되어 있다. 노동공급변수는 경제활동참가이다. (1), (2)열은 자녀의 수가 0명에서 1명으로 올라가는 표본, (3), (4)열은 자녀의 수가 1명에서 2명으로 올라가는 표본, (5), (6)열은 자녀의 수가 2명에서 3명으로 올라가는 표본을 나타낸다. 그리고 각각의 표본에 대하여 홀수 열은 OLS 추정치, 짝수 열은 IV 추정치를 나타낸다.

Table 7: 선행연구와 비교 가능한 자녀의 출산 차이: 전업직장

	(1) OLS 0명에서 1명 출산	(2) IV 출산	(3) OLS 1명에서 2명 출산	(4) IV 출산	(5) OLS 2명에서 3명 출산	(6) IV 출산
자녀의 수	-0.0568*** (0.00409)	0.00905 (0.0151)	-0.0494*** (0.00453)	-0.00279 (0.0165)	-0.0400*** (0.00536)	-0.0334* (0.0183)
교육연수	0.109*** (0.0145)	0.132*** (0.0155)	0.0898*** (0.0163)	0.107*** (0.0174)	0.0912*** (0.0204)	0.0933*** (0.0211)
교육연수의 제곱	-0.00309*** (0.000538)	-0.00384*** (0.000568)	-0.00246*** (0.000608)	-0.00300*** (0.000637)	-0.00249*** (0.000767)	-0.00255*** (0.000782)
연령	0.0595*** (0.00594)	0.0336*** (0.00831)	0.0550*** (0.00775)	0.0394*** (0.00943)	0.0511*** (0.0111)	0.0491*** (0.0123)
연령의 제곱	-0.000736*** (0.0000910)	-0.000411*** (0.000117)	-0.000649*** (0.000116)	-0.000449*** (0.000135)	-0.000598*** (0.000163)	-0.000572*** (0.000177)
임신 당시 연령	-0.0132** (0.00654)	0.000797 (0.00730)	-0.0171** (0.00787)	-0.0102 (0.00824)	-0.0142 (0.0107)	-0.0138 (0.0107)
임신 당시 연령의 제곱	0.0000805 (0.000133)	-0.0000682 (0.000139)	0.000135 (0.000167)	0.0000618 (0.000169)	0.0000556 (0.000235)	0.0000548 (0.000234)
인종: 흑인여부	0.0344*** (0.0106)	0.0386*** (0.0107)	0.0387*** (0.0121)	0.0389*** (0.0121)	0.0408*** (0.0153)	0.0404*** (0.0153)
도시거주 여부	-0.0186** (0.00900)	-0.0206** (0.00909)	-0.00283 (0.0102)	-0.00541 (0.0103)	-0.00347 (0.0130)	-0.00403 (0.0130)
연도 더미	통제	통제	통제	통제	통제	통제
1차 회귀의 F 값		20.16		16.11		11.92
Wu-Hausman p-값		0.00000505		0.00314		0.709
Sargan test p-값		0.254		0.302		0.140
총 관측치 수	12382	12382	9653	9653	5919	5919

주석 표준오차 값은 괄호 안에 표시되어 있다. 노동공급변수는 전업직장 근무 여부이다. (1), (2) 열은 자녀의 수가 0명에서 1명으로 올라가는 표본, (3), (4) 열은 는 자녀의 수가 1명에서 2명으로 올라가는 표본, (5), (6) 열은 자녀의 수가 2명에서 3명으로 올라가는 표본을 나타낸다. 그리고 각각의 표본에 대하여 홀수 열은 OLS 추정치, 짝수 열은 IV 추정치를 나타낸다.

Table 8: 선행연구와 비교 가능한 자녀의 출산 차이: 모든직장

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
자녀의 수	-0.0575*** (0.00397)	0.00485 (0.0147)	-0.0524*** (0.00445)	0.0138 (0.0163)	-0.0430*** (0.00537)	-0.0195 (0.0184)
교육연수	0.128*** (0.0141)	0.150*** (0.0151)	0.104*** (0.0161)	0.129*** (0.0172)	0.0981*** (0.0204)	0.106*** (0.0212)
교육연수의 제곱	-0.00345*** (0.000522)	-0.00416*** (0.000551)	-0.00263*** (0.000598)	-0.00339*** (0.000630)	-0.00235*** (0.000768)	-0.00256*** (0.000784)
연령	0.0569*** (0.00577)	0.0323*** (0.00806)	0.0527*** (0.00762)	0.0305*** (0.00933)	0.0545*** (0.0111)	0.0474*** (0.0123)
연령의 제곱	-0.000707*** (0.0000884)	-0.000400*** (0.000113)	-0.000622*** (0.000114)	-0.000337** (0.000134)	-0.000638*** (0.000164)	-0.000547*** (0.000177)
임신 당시 연령	-0.00945 (0.00635)	0.00378 (0.00708)	-0.0227*** (0.00774)	-0.0130 (0.00815)	-0.0229** (0.0107)	-0.0215** (0.0107)
임신 당시 연령의 제곱	0.0000429 (0.000129)	-0.0000980 (0.000134)	0.000298* (0.000164)	0.000193 (0.000167)	0.000296 (0.000235)	0.000293 (0.000235)
인종: 흑인여부	0.0179* (0.0103)	0.0218** (0.0104)	0.0239** (0.0119)	0.0243** (0.0120)	0.0346** (0.0153)	0.0331** (0.0153)
도시거주 여부	-0.00998 (0.00873)	-0.0118 (0.00882)	-0.00299 (0.0101)	-0.00665 (0.0102)	-0.000689 (0.0130)	-0.00271 (0.0131)
연도 더미	통제	통제	통제	통제	통제	통제
1차 회귀분석의 F 값		20.16		16.11		11.92
Wu-Hausman 검정 p-값		0.00000848		0.0000197		0.180
Sargan 검정 p-값		0.288		0.424		0.114
총 관측치 수	12382	12382	9653	9653	5919	5919

주석: 표준오차 값은 괄호 안에 표시되어 있다. 노동공급변수는 모든 직장 근무 여부이다. (1),(2)열은 자녀의 수가 0명에서 1명으로 올라가는 표본, (3), (4)열은 는 자녀의 수가 1명에서 2명으로 올라가는 표본, (5), (6)열은 자녀의 수가 2명에서 3명으로 올라가는 표본을 나타낸다. 그리고 각각의 표본에 대하여 홀수 열은 OLS 추정치, 짝수 열은 IV 추정치를 나타낸다.

4.2. 가계의 자원

일반적으로 여성의 노동공급 모형에서 남편이 벌어들인 소득을 통제함으로써 비노동 소득이 노동공급에 미치는 효과를 제거한다. 하지만 NSFG에는 가계의 전체소득을 15개의 구간으로 나누어서만 보고하고 있다. 따라서 통상적인 방식으로는 가계의 자원이 여성의 노동공급에 미치는 영향을 통제하기 어렵다.

여기서는 이러한 문제를 극복하기 위해 다음과 같은 전략을 사용한다. 먼저 혼인상태의 여성과 그렇지 않은 여성²을 비교한다. 이는 일차적으로 비노동소득의 유무에 따른 노동공급 효과의 변화를 관찰할 수 있게 한다. 그리고 남편의 학력이 증가함에 따라 남편의 소득, 즉 아내의 입장에서는 비노동소득이 증가할 것을 예상할 수 있다. 따라서 남편의 학력을 비노동소득의 대리변수로 삼아 여성의 노동공급 효과를 관찰한다.

표 9는 혼인상태의 여성과 미혼상태의 여성에 대해서 IV 추정을 한 결과를 나타낸다. 홀수 열은 기혼여성이 표본이며 짝수 열은 미혼여성이 표본이다. 그리고 (1), (2) 열은 경제활동참가율, (3), (4) 열은 전업직장, (5), (6)은 모든직장을 노동공급 변수로 설정하였다. 먼저 기혼여성은 모든 노동공급 변수에서 자녀의 수가 증가할 때 미혼여성에 비하여 노동공급을 줄이는 것으로 나타났다. 기혼여성은 자녀의 수가 증가함에 따라 경제활동참가 확률은 5.02% 하락하고, 전업직장 노동공급 확률은 7.07%, 모든직장의 노동공급 확률은 4.06% 하락하는 것으로 나타난다. 반면 미혼여성의 경우에는 전업직장 노동공급 확률이 3.87% 하락하는 것을 제외하고는 계수가 유의하지 않은 것을 알 수 있다.

그러나 이러한 관찰은 조심스럽게 받아들여져야 한다. 예컨대 기혼여성의 노동공급을 경제활동참가로 측정할 때, 비록 Sargan 검정에서 도구변수는 오차항과

²여기서는 혼인한 적이 없는 여성만을 고려한다. 이혼, 별거, 사별 등은 우리의 고려 대상이 아니다.

Table 9: 가계의 자원 I - 기혼여성과 미혼여성 차이

	(1) 경제활동참가	(2)	(3)	(4) 전업직장	(5)	(6) 모든직장
자녀의 수	-0.0502** (0.0255)	0.0000209 (0.0201)	-0.0707** (0.0269)	0.0387* (0.0211)	-0.0406 (0.0261)	0.00631 (0.0211)
교육연수	0.122*** (0.0207)	0.144*** (0.0278)	0.0928*** (0.0218)	0.0817*** (0.0292)	0.116*** (0.0212)	0.133*** (0.0291)
교육연수의 제곱	-0.00329*** (0.000739)	-0.00372*** (0.00108)	-0.00257*** (0.000778)	-0.00132 (0.00114)	-0.00307*** (0.000757)	-0.00299*** (0.00113)
연령	0.0419*** (0.0140)	0.0540*** (0.0122)	0.0546*** (0.0147)	0.0406*** (0.0128)	0.0449*** (0.0143)	0.0512*** (0.0128)
연령의 제곱	-0.000450** (0.000192)	-0.000838*** (0.000182)	-0.000604*** (0.000202)	-0.000619*** (0.000191)	-0.000496** (0.000196)	-0.000761*** (0.000191)
첫번째 임신 당시 연령	-0.0212** (0.00967)	0.0107 (0.0137)	-0.0308*** (0.0102)	0.0179 (0.0144)	-0.0126 (0.00990)	0.00789 (0.0144)
첫번째 임신 당시 연령의 제곱	0.000258 (0.000178)	-0.000204 (0.000280)	0.000368** (0.000187)	-0.000269 (0.000295)	0.000129 (0.000182)	-0.000127 (0.000294)
인종: 흑인 여부	0.0797*** (0.0199)	0.0281* (0.0144)	0.0913*** (0.0210)	-0.00856 (0.0152)	0.0705*** (0.0204)	-0.0176 (0.0152)
도시 거주 여부	-0.000480 (0.0131)	-0.0136 (0.0139)	-0.00210 (0.0138)	-0.0317** (0.0147)	-0.000604 (0.0134)	-0.0221 (0.0146)
연도터미						
1차 회귀분석 F-값	7.271	10.93	7.271	10.93	7.271	10.93
Wu-Hausman 검정 p-값	0.309	0.00847	0.781	0.0000302	0.194	0.00465
Sargan 검정 p-값	0.194	0.124	0.570	0.800	0.116	0.911
관측치 수	5634	4348	5634	4348	5634	4348

주석 표준오차 값은 괄호 안에 표시되어 있다. 홀수열은 기혼여성, 짝수열은 미혼여성에 대한 추정결과이다. (1), (2) 열은 경제활동참가, (3), (4) 열은 전업직장, (5), (6)은 모든직장을 각각 노동공급 변수로 가진다.

체계적인 관계를 가지지 않지만 1단계 회귀분석의 F-값은 7.27로 기준인 10에 못 미친다. 또한 Wu-Hausman 검정의 p-값이 높아 자녀의 수가 외생적이라는 귀무가설을 기각할 수 없다. 이때 도구변수추정은 적절하지 않을 수 있기 때문이다.

이제 남편의 학력을 비노동소득의 대리변수로 사용하여, 비노동소득에 따라 첫째 자녀의 출산이 여성의 노동공급에 어떤 영향을 주는지 살펴보자.³ 먼저 우리는 남편의 학력을 고등학교 졸업 미만, 고등학교 졸업 또는 동등자격, 그리고 고등학교 졸업 이상의 학력으로 구분하였다. 남편의 학력 증가에 따라 그 자신의 노동소득이 증가할 것이 예상되지만 NSFG는 개별 소득 데이터를 제공하지 않기 때문에 검증이 불가능하다. 하지만 Mincer (1958) 등이 지적하듯이 이는 잘 알려진 사실이므로 받아들이고 진행하였다.

표 10에서 패널 A는 남편의 학력에 따라 기혼여성이 노동공급을 자녀의 숫자가 증가함에 따라 어떻게 조정하는지를 보여준다. 패널 A에서 남편의 학력이 대학 이상일 경우에 경제활동참가, 전업직장, 그리고 모든직장의 노동공급이 감소하는 것을 확인할 수 있다. 경제활동참가 확률은 5.9% 하락하지만 전업직장 노동공급 확률은 9.1% 하락하며, 모든직장 노동공급 확률은 5.7% 하락하는 것을 볼 수 있다. 반면 남편의 학력이 고졸인 경우 여성의 전업직장 노동공급 확률이 5.7% 하락하는 것을 제외하면, 다른 남편의 학력을 가진 경우 자녀의 숫자가 노동공급이 변하지 않는다는 가설을 기각할 수 없었다. 이는 표준적인 노동공급 모형의 예측에 부합하는 것이라고 할 수 있다. 기혼여성의 경우 남편의 소득이 증가할 때 소득효과에 의해 노동공급을 줄일 것으로 예측되기 때문이다.

한편 패널 B에서는 아내의 학력과 남편의 학력이 같은 경우를 보여주고 있다. 그 중에서도 부부의 학력이 대학 이상인 경우에만 경제활동참가는 7.0%, 전업직장은

³이하의 모형표기는 연도더미만을 통제한 최소한의 것이다. 통제변수를 추가하면 우리가 관심을 가지는 계수는 유의해지지 않는 것을 발견하였다. 따라서 소절의 내용은 일반화될 수 없다.

Table 10: 가계의 자원 II - 대리변수로서의 남편의 학력

		경제활동참가	전업직장	모든직장	관측치 수
패널 A					
남편의 학력: 고졸 미만		-0.021 (0.044)	-0.007 (0.042)	0.033 (0.045)	1215
남편의 학력: 고졸		-0.026 (0.030)	-0.0570* (0.033)	-0.034 (0.031)	2010
남편의 학력: 대학 이상		-0.059* (0.031)	-0.091*** (0.034)	-0.057* (0.032)	3429
패널 B					
아내의 학력 = 남편의 학력	아내의 학력: 고졸 미만	-0.024 (0.055)	-0.014 (0.050)	-0.003 (0.055)	603
	아내의 학력: 고졸	-0.017 (0.053)	-0.085 (0.057)	-0.028 (0.055)	765
	아내의 학력: 대학 이상	-0.070*** (0.033)	-0.103*** (0.037)	-0.070*** (0.034)	2702
패널 C					
아내의 학력 > 남편의 학력	아내의 학력: 고졸 미만				
	아내의 학력: 고졸	0.011 (0.025)	0.030 (0.027)	0.013 (0.027)	2382
	아내의 학력: 대학 이상	-0.005 (0.018)	0.027 (0.022)	0.002 (0.020)	3195
패널 D					
아내의 학력 < 남편의 학력	아내의 학력: 고졸 미만	-0.001 (0.044)	-0.002 (0.040)	-0.021 (0.043)	544
	아내의 학력: 고졸	0.040 (0.069)	-0.040 (0.069)	0.043 (0.071)	574
	아내의 학력: 대학 이상				

10.3%, 모든직장은 7.0%의 노동공급 확률이 하락하는 것을 확인할 수 있었다. 다른 학력 조합에서는 자녀의 숫자에 따라 노동공급이 변하지 않는다는 가설을 기각할

수 없었다.

5. 결론

지금까지 첫 번째 임신의 결과가 유산인 경우를 도구변수로 이용하여 자녀가 여성의 노동공급에 미치는 효과를 살펴보았다. 본문의 결과는 자녀의 숫자가 여성의 노동공급에 영향을 미치지 않는다는 것을 보여준다.

더 나아가 해당 도구변수를 이용하면 선행연구들이 다루고 있는 표본도 다룰 수 있다는 점에 착안하여 얻은 추정치를 가지고 선행연구들과 비교해 보아도 자녀의 숫자는 여성의 노동공급에 영향을 미치지 않는 것으로 보인다.

마지막으로, 본 논문은 비노동소득이 노동공급에 미치는 영향을 살펴보았다. 표준적인 예측과 부합하게 자녀를 출산했을 때 기혼 여성이 미혼 여성에 비해 노동공급을 더 크게 줄이는 것으로 나타났다. 또한 남편의 학력을 비노동소득의 대리변수로 보았을 때, 남편의 학력이 대학 이상인 경우 기혼 여성의 노동공급이 감소하는 것을 보았다.

본 연구에는 한계가 존재한다. 자녀의 수가 여성의 노동공급에 미치는 효과가 0이라는 귀무가설을 대부분의 모형표기에서 기각할 수 없었다. 비록 비노동소득을 고려한 모형표기에서는 남편의 학력이 대학 이상일 때에 노동공급이 감소함을 보일 수 있었지만 대부분의 모형표기에서는 그러지 못했다. 대부분의 모형표기에서 도구변수추정의 사후검정을 통해 도구변수의 사용이 적절함에도 불구하고 유의한 결과를 얻지 못했던 것이다.

하지만 상대적으로 많이 연구되어 있는 자녀와 여성의 노동공급에 대해서 미국의 최신의 데이터를 가지고 분석해보았다는 데서 의의를 찾을 수 있다. 나아가 자녀 출산의 진입효과가 중요한 한국이나 일본을 연구하는데 해당 도구변수를 적용해

볼 수 있으리라 생각된다. 한국의 전국 출산력 및 가족보건복지실태조사나 일본의 National Fertility Survey는 이러한 도구변수로 사용될 수 있는 원자료를 제공하므로 응용 가능할 것으로 예상된다.

References

- Angelov, N. and A. Karimi (2012). Mothers' income recovery after childbearing. Working Paper Series, Center for Labor Studies 2012:19, Uppsala University, Department of Economics.
- Angrist, J. and W. Evans (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review* 88(3), 450–477.
- Angrist, J. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion* (1 ed.). Princeton University Press.
- Bronars, S. and J. Grogger (1994). The economic consequences of unwed motherhood: Using twin births as a natural experiment. *American Economic Review* 84(5), 1141–1156.
- Cruces, G. and S. Galiani (2007). Fertility and female labor supply in Latin America: New causal evidence. *Labour Economics* 14(3), 565–573.
- Goldin, C. (1991). The role of World War II in the rise of women's employment. *American Economic Review* 81(4), 741–756.
- Hirvonen, L. (2009). The effect of children on earnings using exogenous variation in family size: Swedish evidence. Working Paper Series 2/2009, Swedish Institute for Social Research.
- Hotz, V. J., S. W. McElroy, and S. G. Sanders (2005). Teenage childbearing

- and its life cycle consequences exploiting a natural experiment. *Journal of Human Resources* 40(3), 683–715.
- Li, X. (2009). Women’s labor market outcomes and childbearing. *New Haven: Yale University. Unpublished.*
- Lundborg, P., E. Plug, and A. W. Rasmussen (2014). Fertility effects on female labor supply: IV evidence from IVF treatments. IZA Discussion Papers 8609, Institute for the Study of Labor (IZA).
- Martin, J. A., B. E. Hamilton, and M. J. Osterman (2012). Three decades of twin births in the United States, 1980-2009. NCHS data brief no 80, National Center for Health Statistics.
- Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of political economy* 66(4), 281–302.
- Regan, L. and R. Rai (2000). Epidemiology and the medical causes of miscarriage. *Best practice & research Clinical obstetrics & gynaecology* 14(5), 839–854.
- Rosenzweig, M. and K. I. Wolpin (1980). Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica* 48(1), 227–240.
- Schultz, T. (2007). Fertility in developing countries. Working papers, Economic Growth Center, Yale University.

A. 부록

A.1. Hotz et al. (2005)의 추정전략

부록에서는 본문과 다른 추정전략을 통해서 도구변수추정을 수행한다. Hotz et al. (2005)는 여성의 임신의 결과(D)가 출산(B), 낙태(A), 유산(M)으로 나타날 수 있다고 보았다. 한편 여성의 잠재유형(D*)은 유산을 겪지 않았을 때 출산(B*)을 하거나 낙태(A*)를 하는 결정으로 구성된다고 보았다. 저자들은 첫번째 임신에서 출산을 한 여성과 출산을 미룬 여성 사이의 평균효과를 관심의 대상인 인과효과로 정의했다:

$$\beta = E[Y_B - Y_{B*}|D = B]. \quad (3)$$

그런데 우리가 알지 못하는 $E[Y_{B*}|D = B]$ 를 구하기 위해서 유산을 겪는 여성은 그렇지 않았다면 출산을 했을 잠재유형(B*)과 낙태를 했을 잠재유형(A*)으로 구성된다는 가정을 사용할 수 있다. 만약 B* 유형의 비율을 P^* 라고 하면 $E[Y_M]$ 은 다음과 같다:

$$E[Y_M] = P^*E[Y_{B*}] + (1 - P^*)E[Y_{A*}]. \quad (4)$$

식 4 를 $E[Y_{B*}]$ 에 대해서 정리하면 다음을 얻을 수 있다:

$$E[Y_{B*}] = \frac{E[Y_M] - (1 - P^*)E[Y_{A*}]}{P^*}. \quad (5)$$

저자들은 모든 유산이 임의적으로 발생하고 모든 출산력이 정확히 보고될 때, 유산한 사람 중 그렇지 않았다면 출산을 했을 비율 P^* 는 유산하지 않은 사람들

중 출산을 한 여성들의 비율 P 와 같아진다고 주장한다. 더 나아가 유산과 낙태가 평균적으로 종속변수에 동일한 직접적인 영향을 준다면 $E[Y_A] = E[Y_{A*}]$ 라고 주장한다. 이때

$$E[Y_{B*}] = \frac{E[Y_M] - (1 - P)E[Y_A]}{P} \quad (6)$$

이 성립하게 된다. 식 3 과 식 6을 결합해보면 다음과 같다:

$$\begin{aligned} \beta &= E[Y_B - Y_{B*} | D = B] \\ &= \frac{PE[Y_B] + (1 - P)E[Y_A] - E[Y_M]}{P} \\ &= \frac{E[Y_{\tilde{M}} - Y_M]}{P} \\ &= \frac{Cov(Y, M)}{Cov(B, M)}. \end{aligned} \quad (7)$$

따라서 식 7은 낙태를 포함하여 도구변수 추정법을 적용해야 함을 의미한다. 본문에서는 낙태를 표본에서 제외했지만 부록에서는 이를 따라서 본문과 동일한 내용들을 추정해보도록 한다.

A.2. 기본결과

표 11은 노동공급을 경제활동참가로 측정했을 때 자녀의 수가 노동공급에 어떤 영향을 주는지 나타내는 것이다. 1차 회귀분석의 F-값은 가장 낮은 회귀식이 35.3으로 기준인 10을 넘는다. 또한 유의수준 0.1에서 자녀의 수는 내생적이라고 할 수 있으며, 또한 (2)열을 제외하면 도구변수는 유의수준 0.05에서 적절한 것으로 보인다.

(3)열부터 (6)열을 기준으로 OLS 추정치는 대체로 5.7%의 경제활동참가 확률의 하락을 예측하지만 IV 추정치는 3.4%의 하락을 예측한다. OLS 추정치가 계수의 크기를 과대평가하고 있는 것을 확인할 수 있다.

표 12에서는 모든 노동공급 변수를 나타내고 있다. 1차 회귀식에서의 F-값은 기준을 충족시키고, 자녀의 숫자는 내생적인 것으로 나타난다. 도구변수 또한 적절한 것으로 보인다. 결과를 살펴보면, OLS 추정치는 IV 추정치에 비해 계수를 과대추정하는 것을 볼 수 있다. 예컨대 OLS 추정치의 경우 자녀의 수가 늘어날 때 5.83% 전업직장 근무 확률이 낮아질 것으로 예측하지만, IV 추정치는 2.48%만 낮아질 것으로 예측하고 있다. 모든 직장의 경우는 5.66% 근무 확률의 감소를 예측하는 OLS 추정치와 달리, IV 추정치는 2.53% 하락하리라 예측하고 있다.

A.3. 선행연구와 표본

표 13의 목적은 우리의 도구변수를 첫 번째 임신 뿐만 아니라, 두 번째 임신, 세 번째 임신 등으로 확장함으로써 표본을 선행연구의 그것에 대응시키기 위함이다.

경제활동참가를 노동공급 변수로 보는 표 13은 다음과 같이 구성되어 있다. (1), (2)열은 자녀의 수가 0명에서 1명으로 늘어날 수 있는 첫번째 임신을 한 여성들로 구성되어 있다. 이들은 표 3에서 다룬 표본 그리고 Li (2009)와 동일하다. 그리고 (3), (4)열은 자녀의 수가 1명에서 2명으로 늘어날 수 있는 두 번째 임신을 한 여성들로 구성되어 있다. 이들은 Bronars and Grogger (1994)이 다루는 표본에 대응된다. 한편 (5), (6)열은 자녀의 수가 2명에서 3명으로 올라갈 수 있는 세 번째 임신을 한 여성들로 구성되어 있다. 이들은 Angrist and Evans (1998)에 대응된다. 마지막으로 각각의 표본에 대하여 홀수 열은 OLS 추정치, 짝수 열은 IV 추정치를 나타낸다.

자녀가 없다가 1명 생길 때에는 자녀 수가 증가할 때 경제활동참가율이 3.37%

Table 11: OLS vs. IV 추정치 비교: 경제활동참가

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
자녀의 수	-0.0487*** (0.00304)	-0.0639*** (0.00871)	-0.0572*** (0.00344)	-0.0341*** (0.00846)	-0.0570*** (0.00344)	-0.0337*** (0.00846)
교육연수			0.138*** (0.0124)	0.147*** (0.0127)	0.132*** (0.0125)	0.141*** (0.0128)
교육연수의 제공			-0.00404*** (0.000445)	-0.00429*** (0.000452)	-0.00380*** (0.000447)	-0.00406*** (0.000454)
연령			0.0472*** (0.00514)	0.0381*** (0.00599)	0.0473*** (0.00515)	0.0382*** (0.00599)
연령의 제공			-0.000598*** (0.0000784)	-0.000483*** (0.0000875)	-0.000599*** (0.0000785)	-0.000483*** (0.0000875)
첫번째 임신 당시 연령			-0.0105* (0.00553)	-0.00655 (0.00570)	-0.00817 (0.00555)	-0.00412 (0.00572)
첫번째 임신 당시 연령의 제공			0.0000635 (0.000113)	0.0000200 (0.000114)	0.0000274 (0.000113)	-0.0000173 (0.000114)
인종: 흑인 여부					0.0477*** (0.00878)	0.0486*** (0.00877)
도시 거주 여부					-0.00429 (0.00774)	-0.00415 (0.00774)
연도더미	통제	통제	통제	통제	통제	통제
1차 회귀분석 F-값		35.31		53.31		53.20
Wooldridge의 Score p-값 (외생성)		0.0705		0.00357		0.00332
Wooldridge의 Score p-값 (도구변수적절)		0.0000545		0.0750		0.0540
관측치 수	14232	14232	14232	14232	14232	14232

주석: 표준오차 값은 괄호 안에 표시되어 있다. 종속변수는 경제활동참가 여부이다. 홀수 열은 OLS 추정의 결과이며 짝수 열은 IV 추정의 결과로 대칭적으로 비교할 수 있다.

Table 12: OLS vs. IV 추정치 비교: 경제활동참가, 전업직장, 모든직장

	(1) 경제활동참가	(2) 경제활동참가	(3) 전업직장	(4) 전업직장	(5) 모든직장	(6) 모든직장
자녀의 수	-0.0570*** (0.00334)	-0.0337*** (0.00846)	-0.0583*** (0.00362)	-0.0248** (0.00966)	-0.0566*** (0.00349)	-0.0253*** (0.00906)
교육연수	0.132*** (0.0125)	0.141*** (0.0128)	0.0973*** (0.0136)	0.110*** (0.0138)	0.122*** (0.0131)	0.134*** (0.0132)
교육연수의 제곱	-0.00380*** (0.000459)	-0.00406*** (0.000454)	-0.00269*** (0.000498)	-0.00306*** (0.000504)	-0.00334*** (0.000480)	-0.00368*** (0.000472)
연령	0.0473*** (0.00503)	0.0382*** (0.00599)	0.0626*** (0.00546)	0.0495*** (0.00618)	0.0547*** (0.00526)	0.0424*** (0.00614)
연령의 제곱	-0.000599*** (0.0000773)	-0.000483*** (0.0000875)	-0.000792*** (0.0000839)	-0.000626*** (0.0000907)	-0.000687*** (0.0000808)	-0.000531*** (0.0000897)
첫번째 임신 당시 연령	-0.00817 (0.00558)	-0.00412 (0.00572)	-0.00689 (0.00606)	-0.00107 (0.00635)	-0.00365 (0.00583)	0.00180 (0.00598)
첫번째 임신 당시 연령의 제곱	0.0000274 (0.000115)	-0.0000173 (0.000114)	-0.0000260 (0.000125)	-0.0000901 (0.000128)	-0.0000432 (0.000120)	-0.000103 (0.000120)
인증: 흑인 여부	0.0477*** (0.00900)	0.0486*** (0.00877)	0.0348*** (0.00978)	0.0361*** (0.00976)	0.0174* (0.00942)	0.0186** (0.00936)
도시 거주 여부	-0.00429 (0.00771)	-0.00415 (0.00774)	-0.0156* (0.00837)	-0.0154* (0.00839)	-0.0121 (0.00806)	-0.0119 (0.00812)
연도더미	통제	통제	통제	통제	통제	통제
1차 회귀분석 F-값		53.20		53.20		53.20
Wooldridge의 Score p-값 (외생성)		0.00332		0.000158		0.000196
Wooldridge의 Score p-값 (도구변수 적절)		0.0540		0.116		0.445
관측치 수	14232	14232	14232	14232	14232	14232

주석: 표준오차 값은 괄호 안에 표시되어 있다. 종속변수는 (1), (2)열은 경제활동참가 여부이고 (3), (4)열은 전업직장, (5), (6)열은 모든직장이다. 홀수 열은 OLS 추정의 결과이며 짝수 열은 IV 추정의 결과이다.

감소한다. 그리고 자녀가 1명에서 2명으로 늘어날 때에는 자녀 수가 증가할 때 경제활동참가율이 2.89% 감소한다. 마지막으로 자녀가 2명에서 3명으로 늘어날 때에는 자녀 수가 증가할 때 경제활동참가율이 4.44% 감소하는 것으로 나타난다.

A.4. 가계의 자원

표 14는 혼인상태의 여성과 미혼상태의 여성에 대해서 IV 추정을 한 결과를 나타낸다. 홀수 열은 기혼여성이 표본이며 짝수 열은 미혼여성이 표본이다. 그리고 (1), (2) 열은 경제활동참가율, (3), (4) 열은 전업직장, (5), (6)은 모든직장을 노동공급 변수로 설정하였다. 자녀의 수에 대한 경제활동참가 여부와 모든직장 근로 여부 등 노동공급의 감소가 기혼여성이 미혼여성에 비하여 크다는 것을 알 수 있다. 경제활동참가율은 기혼여성은 4.07% 하락이 예측되지만 미혼여성은 3.36% 하락이 예측된다. 모든 직장 근로 확률의 경우 기혼여성은 3.67% 하락이 예측되지만 미혼여성은 3.05% 하락이 예측되기 때문이다.

표 15는 비노동소득의 대리변수로 남편의 학력을 살펴본다. 패널 A에서 남편의 학력이 고졸 미만일 때에는 자녀의 수가 증가할 때 노동공급을 증가한다는 증거가 없다. 하지만 남편의 학력이 고졸일 때는 자녀의 수가 증가할 때 7.64% 경제활동참가율이 떨어지며, 남편의 학력이 대학 이상일 때는 8.51%로 고졸 남편을 가진 여성에 비해 노동공급이 크게 감소한다. 이러한 패턴은 전업직장과 모든직장을 노동공급 변수로 측정했을 때에도 동일하게 나타나는 것으로 보인다. 이러한 현상은 비노동소득이 증가함에 따라 소득효과에 의해 노동이 감소할 것이라는 표준적인 노동공급의 예측과 부합하는 것으로 보인다.

한편 아내와 남편의 학력이 동등하면서 둘 다 대학 이상의 학력을 가진 기혼여성일 경우, 자녀의 수가 증가함에 따라 경제활동참가율이 8.96%, 전업직장 근무 확률이 12.86%, 모든 직장 근무 확률이 8.81% 하락하는 것을 볼 수 있다.

Table 13: 선행연구와 비교 가능한 자녀의 출산 차이: 경제활동참가

	(1) OLS 0명에서 1명 출산	(2) IV 1명에서 2명 출산	(3) OLS 1명에서 2명 출산	(4) IV 2명에서 3명 출산	(5) OLS 2명에서 3명 출산	(6) IV 3명에서 4명 출산
자녀의 수	-0.0570*** (0.00334)	-0.0337*** (0.00846)	-0.0507*** (0.00397)	-0.0289** (0.0114)	-0.0468*** (0.00484)	-0.0444*** (0.0141)
교육연수	0.132*** (0.0125)	0.141*** (0.0128)	0.111*** (0.0147)	0.119*** (0.0150)	0.108*** (0.0190)	0.109*** (0.0193)
교육연수의 제곱	-0.00380*** (0.000459)	-0.00406*** (0.000454)	-0.00301*** (0.000547)	-0.00325*** (0.000540)	-0.00289*** (0.000716)	-0.00292*** (0.000703)
연령	0.0473*** (0.00503)	0.0382*** (0.00599)	0.0337*** (0.00691)	0.0262*** (0.00792)	0.0386*** (0.0100)	0.0378*** (0.0110)
연령의 제곱	-0.000599*** (0.0000773)	-0.000483*** (0.0000875)	-0.000384*** (0.000104)	-0.000288** (0.000115)	-0.000454*** (0.000148)	-0.000444*** (0.000159)
첫번째 임신 당시 연령	-0.00817 (0.00558)	-0.00412 (0.00572)	-0.0149** (0.00717)	-0.0118 (0.00731)	-0.0200** (0.00993)	-0.0198** (0.00965)
첫번째 임신 당시 연령의 제곱	0.0000274 (0.000115)	-0.0000173 (0.000114)	0.000133 (0.000152)	0.0000965 (0.000152)	0.000208 (0.000220)	0.000207 (0.000210)
인종: 흑인 여부	0.0477*** (0.00900)	0.0486*** (0.00877)	0.0524*** (0.0107)	0.0528*** (0.0105)	0.0520*** (0.0137)	0.0519*** (0.0134)
도시 거주 여부	-0.00429 (0.00771)	-0.00415 (0.00774)	0.000420 (0.00921)	-0.000392 (0.00925)	0.00203 (0.0119)	0.00187 (0.0120)
연도터미	통제	통제	통제	통제	통제	통제
1차 회귀분석 F-값		53.20		41.22		44.41
Wooldridge의 Score p-값 (외생성)		0.00332		0.0420		0.857
Wooldridge의 Score p-값 (도구변수 적절)		0.0540		0.237		0.213
관측치 수	14232	14232	10553	10553	6552	6552

주석 표준오차 값은 괄호 안에 표시되어 있다. 노동공급변수는 경제활동참가이다. (1), (2)열은 자녀의 수가 0명에서 1명으로 올라가는 표본, (3), (4)열은 자녀의 수가 1명에서 2명으로 올라가는 표본, (5), (6)열은 자녀의 수가 2명에서 3명으로 올라가는 표본을 나타낸다. 그리고 각각의 표본에 대하여 홀수 열은 OLS 추정치, 짝수 열은 IV 추정치를 나타낸다.

Table 14: 가계의 자원 I - 기혼여성과 미혼여성 차이

	(1)	(2)	(3)	(4)	(5)	(6)
	경제활동참가			전업직장		모든직장
자녀의 수	-0.0407** (0.0161)	-0.0336*** (0.0125)	-0.0460** (0.0182)	-0.0161 (0.0139)	-0.0367** (0.0169)	-0.0305** (0.0135)
교육연수	0.126*** (0.0190)	0.146*** (0.0218)	0.0950*** (0.0198)	0.0650*** (0.0243)	0.118*** (0.0193)	0.132*** (0.0231)
교육연수의 제곱	-0.00348*** (0.000664)	-0.00409*** (0.000796)	-0.00268*** (0.000709)	-0.000956 (0.000930)	-0.00317*** (0.000679)	-0.00334*** (0.000858)
연령	0.0339*** (0.0115)	0.0608*** (0.00942)	0.0482*** (0.0118)	0.0650*** (0.00959)	0.0414*** (0.0117)	0.0625*** (0.00974)
연령의 제곱	-0.000347** (0.000163)	-0.000903*** (0.000146)	-0.000547*** (0.000168)	-0.000933*** (0.000149)	-0.000455*** (0.000166)	-0.000893*** (0.000151)
첫번째 임신 당시 연령	-0.0177** (0.00833)	0.0100 (0.0113)	-0.0250*** (0.00908)	0.0108 (0.0121)	-0.0131 (0.00860)	0.00888 (0.0120)
첫번째 임신 당시 연령의 제곱	0.000223 (0.000162)	-0.000232 (0.000238)	0.000320* (0.000178)	-0.000230 (0.000258)	0.000156 (0.000168)	-0.000203 (0.000253)
인종: 흑인 여부	0.0724*** (0.0173)	0.0289** (0.0127)	0.0890*** (0.0198)	-0.0000392 (0.0136)	0.0600*** (0.0181)	-0.00768 (0.0135)
도시 거주 여부	0.00208 (0.0124)	-0.0181 (0.0123)	0.00281 (0.0132)	-0.0293** (0.0132)	0.00435 (0.0128)	-0.0257** (0.0130)
연도터미	통제	통제	통제	통제	통제	통제
1차 회귀분석 F-값	46.29	46.02	46.29	46.02	46.29	46.02
Wooldridge의 Score p-값 (외생성)	0.0321	0.153	0.0974	0.00894	0.0352	0.0964
Wooldridge의 Score p-값 (도구변수 적절)	0.471	0.0112	0.726	0.476	0.678	0.870

관측치 수 6222 5276 6222 5276 6222 5276

주석 표준오차 값은 괄호 안에 표시되어 있다. 홀수열은 기혼여성, 짝수열은 미혼여성에 대한 추정결과이다. (1), (2) 열은 경제활동참가, (3), (4) 열은 전업직장, (5), (6)은 모든직장을 각각 노동공급 변수로 가진다.

Table 15: 가계의 자원 II - 대리변수로서의 남편의 학력

		경제활동참가	전업직장	모든직장	관측치
패널 A					
남편의 학력: 고졸 미만		-0.0558 (0.0466)	-0.0077 (0.0442)	0.0001 (0.0472)	980
남편의 학력: 고졸		-0.0764** (0.0317)	-0.0884** (0.0346)	-0.0767** (0.0328)	1705
남편의 학력: 대학 이상		-0.0851*** (0.0265)	-0.1260*** (0.0304)	-0.0854*** (0.0277)	3528
패널 B					
아내의 학력 = 남편의 학력	남편의 학력: 고졸 미만	-0.0458 (0.0568)	-0.0234 (0.0498)	-0.0290 (0.0581)	484
	남편의 학력: 고졸	-0.0485 (0.0526)	-0.0604 (0.0599)	-0.0382 (0.0539)	649
	남편의 학력: 대학 이상	-0.0896*** (0.0290)	-0.1286*** (0.0335)	-0.0881*** (0.0302)	2817
패널 C					
아내의 학력 > 남편의 학력	남편의 학력: 고졸 미만				
	남편의 학력: 고졸	0.0116 (0.0526)	0.0462 (0.0502)	0.0399 (0.0528)	317
	남편의 학력: 대학 이상	-0.0511 (0.0350)	-0.0861*** (0.0423)	-0.0126 (0.0381)	928
패널 D					
아내의 학력 < 남편의 학력	남편의 학력: 고졸 미만	-0.0717 (0.0459)	-0.0668 (0.0429)	-0.0929*** (0.0454)	458
	남편의 학력: 고졸	-0.0141 (0.0578)	-0.0594 (0.0584)	-0.0147 (0.0597)	566
	남편의 학력: 대학 이상				

Children and Female Labor Supply :Using the Miscarriage as an Instrumental Variable

Kyusik Kum

Seoul National University

February 2018

Abstract

In this paper, we analyze how the number of children affects the female labor supply. Using the miscarriage in the first pregnancy known for its random occurrence in the medical science as an instrumental variable, we estimate the effect of the number of children on the female labor supply. Our results based on the most recent data of *National Survey of Family Growth* in various specifications show that the female labor supply may not be affected by the number of children. .

Keyword: labor supply, natural experiment, miscarriage, abortion, instrumental variable

Student number: 2012-20148.